

Ethics for Experimental Manipulation of Religion

Richard A. Nielsen*

December 2, 2014

Abstract

This chapter makes a distinction between *measurement manipulations* and *change manipulations* in experiments involving religion. Measurement manipulations are experimental interventions that allow researchers to measure some aspect of religiosity or some characteristic of a religious individual. Change manipulations are experimental interventions in which researchers attempt to set some aspect of a subject's religiosity to a level or state that it would otherwise not attain. I argue that experiments that look more like measurement manipulations will generally be ethical, while those that look more like change manipulations are less likely to be ethical. I illustrate with hypothetical and real examples. The conclusions suggest that it will be difficult to learn causal knowledge about the effects of religion from experiments because there are serious ethical problems with setting the religiosity of subjects to levels that the subjects do not choose.

*Assistant Professor, Department of Political Science, MIT. Thanks to Scott Desposato for hosting the Ethics in Comparative Politics Experiments Conference at UCSD on May 1-2, 2013. Thanks to Scott Desposato, Becky Morton, Lucas Stanczyk, Peter McMurray, Jesse Driscoll, Chana Teeger, Adria Lawrence, Gina Bateson, and Emma Smith for helpful discussions.

In this chapter, I explore the ethical boundaries of social science experiments that involve religion. I argue for a distinction: experimental manipulations that allow *measurement* of an individual's religious beliefs, practice, or experience are generally ethical, while manipulations that *change* an individual's religious beliefs, practice, or experience are more likely to be unethical.

Is it ethical to change the religiosity of experimental subjects to learn how religiosity affects political attitudes and behaviors? This question is crucially important to the field of religion and politics, and Comparative Politics more generally. As the discipline of Political Science makes an experimental turn, fields where ethical experiments are infeasible will be at a structural disadvantage. Knowledge production within these fields is still as important as before; there is no reason to believe that the ability to ethically design experimental manipulations is correlated with the substantive importance of a given field. However, the inability to do ethical experiments may mean that (1) fewer researchers work on religion and politics because the field does not ethically allow certain approaches, and (2) it is harder to create new knowledge about the causal effects of religion. My conclusion is that meaningful interventions that change religiosity are likely to be unethical, corroborating earlier debates on psychology (Batson, 1977; Yeatts and Asher, 1979; Batson, 1979).

Given the sensitive nature of religious experience (one of the topics not discussed in polite conversation!), I feel that I should briefly describe my religious commitments. I was born and raised a Mormon (the Church of Jesus Christ of Latter-day Saints, or LDS church), attended Brigham Young University, and spent two years as a Mormon missionary in Alaska and Canada. I am currently practicing, but not believing, meaning that I participate in the Mormon community and Mormon religious rituals, but do not believe the foundational truth-claims of the religion. This background gives me a relatively rich ethnographic understanding of the experience of religious faith, the concerns of at least some segment of religious people, and the costs and benefits associated with changes in religiosity. Throughout, I appeal to religious examples. Because of my background, I am most confident in the examples based on Mormonism.

I begin in Section 1 by presenting a distinction between two ideal types of experiments involving religion: those that *measure* religious beliefs and those that *change* them. I argue that measurement experiments will generally be ethical (with respect to issues of religion) while experiments that

change religiosity are less likely to be ethical. Section 2 provides three arguments to support my assertion that change manipulations are often unethical: (1) because individuals have a strong interest in forming beliefs about questions of ultimate value in non-manipulative circumstances, it is difficult to experimentally change religiosity without imposing a possible harm; (2) because religion makes untestable claims about harms and benefits, it is impossible to change religiosity without imposing a possible harm; and (3) that it is impossible to provide adequate informed consent about changing subjects' religion without undermining the inferential value of the study. Section 3 discusses an experiment I proposed that raised the issues discussed here as well as evaluations of other experimental studies according to the criteria I suggest.

1 Measurement and Change Manipulations

I argue that experiments involving religion fall on a spectrum between two ideal types. The first ideal type uses experimental manipulation to elicit and measure religious beliefs, practices, or experience. The second ideal type uses experimental manipulation to change religious beliefs, practices, or experience. For ease of discussion, I refer to these respectively as *measurement manipulations* and *change manipulations*.¹ I draw a fundamental distinction between these two ideal types and suggest that change manipulations face ethical challenges that measurement manipulations do not. Figure 1 shows these two ideal types as ends of a spectrum with my placement of some possible experimental designs along the spectrum.

Measurement manipulations are those that use experimental manipulation to measure the level of some attribute of an individual, group, or environment. Table 1 includes examples of this type of manipulation. The key feature of this type of manipulation is that the researcher is not attempting to assign the subject to a particular level of variable. Rather, the researcher is attempting to assess the “natural” level of a particular variable as unobtrusively as possible. This is not to say that these experiments do not change or intrude on subjects. Results from fields as disparate as anthropology

¹I define experimental manipulation as a change to an individual, group, or environment as a result of actions by a scientific researcher. My definition has an embedded causal assumption: that the change would not have occurred without actions by a researcher. A second key distinction is that the change results from actions of scientific researchers, rather than individuals of other backgrounds.

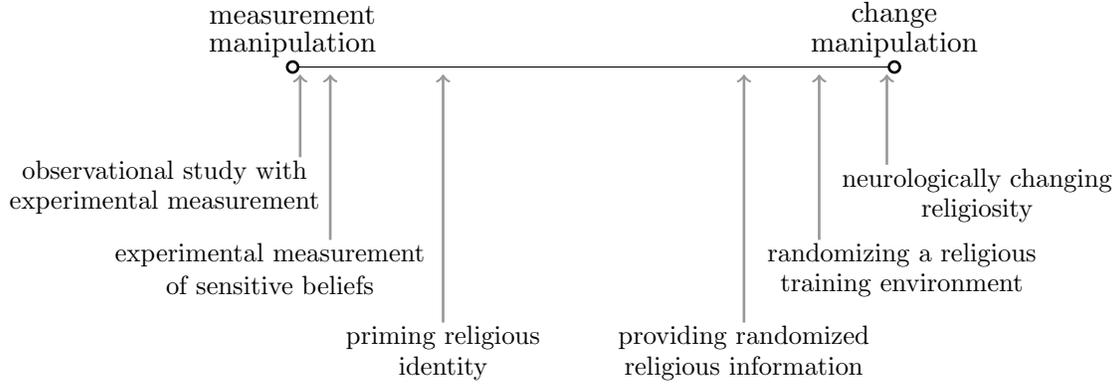


Figure 1: A depiction of the spectrum from measurement manipulations to change manipulations with my subjective assessment of where some possible experimental designs fall on this spectrum.

and physics suggest that the act of observing changes the thing that is observed. However, the intent of measurement manipulations is to produce measurements that are maximally accurate and minimally intrusive. The results of measurement manipulations are desirable because they can reveal attitudes and behaviors that subjects are unwilling or unable to express directly in a survey or interview.

Table 1: Examples of Measurement Manipulations

-
- (1) Researchers randomize the contents of a survey to elicit responses to sensitive questions. (List Experiment)
 - (2) In a survey setting, researchers randomize which individuals or groups are reported to support particular policies or actions. (Endorsement Experiment).
 - (3) Researchers randomize elements of a fictional story to measure differences in responses.
 - (4) Researchers use of dictator games with confederates to assess altruism.
-

For example, in a list experiment attempting to measure racism in the American South (Kuklinski, Cobb and Gilens, 1997), the researchers used an experimental manipulation to allow respondents to be honest about racism without revealing that racism to the survey administrator. Although there was an experimental manipulation, the intent was to be *less* invasive and more accurate than directly asking “Are you a racist?”

It is possible that a measurement manipulation will change subject religiosity. In fact, because of the possibility that measurement inherently changes the thing that is measured, there may be no pure measurement manipulations. Even manipulations intended to minimize the intrusiveness of measurement may still induce changes. In the example of a list experiment on racism, the fact that a survey researcher has bothered to ask about racism may create new thoughts and attitudes in individuals who respond to the list experiment. Even if the researcher cannot tell what answers a respondent gives, the act of acknowledging racism in some domain could either decrease or increase feelings of racism. It is impossible to know whether measurement induces these types of changes, because this would require another measurement which could also induce more changes. Thus, the closest we can confidently come to this logical end of the spectrum is that some experiments are primarily measurement manipulations.

Change manipulations do not simply measure some aspect of a trait or characteristic. Rather, they are designed to set that trait at a particular value so that the effect of the trait can be observed. This process of setting subject characteristics via experimental manipulation results in changes to the subject by definition. Table 2 gives examples of experimental designs with change manipulations.

Table 2: Examples of Change Manipulations

(1)	Telling subjects new information that affects their beliefs and attitudes.
(2)	Manipulating a religious environment or experience to induce changes in subjects' beliefs and attitudes.
(3)	Encouraging subject participation in a religious ceremony suspected to change religious commitment or identity.

In the physical sciences, it is possible for researchers to directly manipulate features of the object of study to estimate the effects of those features. Social scientists studying organizations can directly change aspects of those organizations (say, removing network ties or changing organizational structure). However, researchers who study individuals cannot generally directly set the beliefs or attitudes of individuals at a desired level. In the context of religion, this means that there is currently no way for a researcher to fix the religiosity of an individual at some level. Instead,

researchers attempt to manipulate the beliefs and attitudes of humans by providing them with new information, by withholding information, by constraining a subject's environment, or by triggering some physical process that researchers believe is associated with an attitudinal change.

This means that there are no manipulations of religiosity that are purely change manipulations according to my definition. Researchers cannot be certain that they have set religiosity to a particular level. Instead, they can only randomize variables that affect religiosity and identify intent-to-treat effects of those interventions or use compliance information to estimate the effects of changes of religiosity in an instrumental variables set-up (where the randomized variable serves as an instrument for changes in religiosity).

In experiments on religion that identify intent-to-treat effects, the researcher can be thought of as testing the effects of an intervention that has religious side-effects. This type of research does not demonstrate any causal effect of religion without assumptions about compliance. Alternatively, experiments that use the manipulation as an instrumental variable for religiosity will necessarily have to measure religiosity following any manipulation. This means that pure change manipulations on religious beliefs and attitudes are not possible because the researcher cannot be confident whether religiosity was manipulated or the measurement has error. All this is a result of the inability of researchers to set the attitudes or beliefs of an individual to a specified level or category.

Thus, experiments on religion fall on a spectrum between pure measurement manipulations and pure change manipulations.

2 Change Manipulations of Religiosity are Generally Unethical

My discussion of the ethics of experiments involving religion follows from my distinction between measurement and change manipulations. I argue that a pure measurement manipulation of religion is generally ethical unless it violates some other ethical principle of experimentation such as deception or harm to the subject. On the other hand, pure change manipulations of religion are unethical, even if other aspects of the experiment are ethical (subjects are fully informed, etc). Because all experimental manipulations fall between these ideal types, I argue that the more an experiment approximates a change manipulation, the less ethically defensible it is likely to be.

2.1 An Argument from Interest in Forming Beliefs Without Manipulation

- (a) People have a strong interest in forming beliefs, especially beliefs of ultimate value.
- (b) People have a strong interest in forming these beliefs in favorable, non-manipulative circumstances.
- (c) Change manipulations require that beliefs be formed under manipulative circumstances.
- (d) Potential harm from such manipulations cannot be fully disclosed prior to manipulation and may not be reversible.
- (e) Therefore the harm should be taken seriously and is often large enough to outweigh benefits of the research.

(a) People have a strong interest in forming beliefs, especially beliefs of ultimate value.

Human beings display strong interest in forming beliefs about themselves, other people, and the nature of humanity's physical and metaphysical surroundings. Beliefs of "ultimate value" are beliefs that deal with the metaphysics of human existence. These include beliefs about deity, the origins and purpose of human existence, whether aspects of an individual persist beyond death, and so on. While not all individuals seem equally driven to form such beliefs, virtually all individuals form beliefs to some degree and most hold these beliefs to be important.

(b) People have a strong interest in forming these beliefs in favorable, non-manipulative circumstances.

Individuals have a strong interest in forming beliefs about ultimate value in non-manipulated environments. Beliefs about "facts" are often passed on in manipulated environments (children learning multiplication tables in a classroom), but it can be problematic for someone to lead other individuals to hold beliefs about ultimate value through manipulation of their environment. Most modern conceptions of human rights enshrine such an interest in terms such as "freedom of thought" or "freedom of expression." This interest might also be perceptible in norms against censorship. While this interest in forming beliefs in non-manipulated environments exists for many types of beliefs, it is especially important for individuals forming beliefs about matters of ultimate value. This

may be seen in the additional rights of “freedom of conscience” and “freedom of religion” that are additional to freedoms of thought and expression.

(c) Change manipulations require that beliefs be formed under manipulative circumstances.

Manipulation is at the essence of experimental assignment of treatment and control. In order to experimentally learn about the effects of religiosity, it must be randomly assigned, and this “assignment” requires manipulation. Thus, change manipulations that set a subject’s religiosity to a value it would not have attained naturally are forming the subject’s beliefs under manipulative circumstances. Any experiment that does not manipulate the subject’s beliefs is, by definition, not a change manipulation. Because these beliefs are about religion, they are likely to involve beliefs about ultimate value. Manipulating such beliefs is very likely to run counter to subjects’ strong interest in forming their beliefs about ultimate value in non-manipulated environments.

(d) Potential harm from such manipulations cannot be fully disclosed prior to manipulation and may not be reversible.

Research in many settings requires subjects to form beliefs in manipulated environments. Most studies involving deception manipulate subjects’ beliefs about the purpose of the research or the nature of the experimental task in order to maintain the validity of the experiment. This manipulation of beliefs is typically justified on the grounds that it can be corrected by debriefing subjects afterward to correct their misperceptions. Alternatively, it is justified by positing that subjects do not have a vital interest in knowing the information about which they were led to form incorrect beliefs. For example, it seems reasonable that subjects generally do not have a strong interest in knowing a researcher’s precise hypothesis while participating in an experiment. Such knowledge can reasonably be withheld until after participation is complete because it is not generally vital to the subject to know and any incorrect beliefs can be easily addressed through debriefing.

Manipulating religious beliefs is problematic because there is potential for harm that cannot necessarily be mitigated through debriefing. Researchers cannot fully comprehend the harms that

might come to respondents whose religiosity is manipulated, so they almost certainly cannot anticipate and mitigate these harms. Moreover, at least some manipulations of religious ideas and beliefs may not be fully reversible. Critical literature cannot be unread, doubt cannot be undoubted, and transcendent experiences are hard to shake, even if researchers subsequently inform subjects that these experiences were manipulated.

(e) Therefore the harm should be taken seriously and is often large enough to outweigh benefits of the research.

If individuals have a strong interest in forming beliefs about ultimate value in non-manipulated environments, then change manipulations of religiosity will be difficult to justify on ethical grounds. There may be societal or scientific interests that trump this interest — it may on balance be ethical to manipulate the beliefs of religious terrorists to encourage pacifism — but such justification will require argument on a case-by-case basis. In general, subjects will rightfully have an aversion to experiments that manipulate their religiosity because of the value they place on forming their own judgments about things of ultimate value.

2.2 An Argument from the Ambiguity of Religious Harms and Benefits

- (a) Religion makes untestable claims about harms and benefits.
- (b) People with different beliefs can view identical treatments as clear harm or clear benefit.
- (c) All treatments that meaningfully change religiosity will be a clear harm to someone.
- (d) Such a claim of harm is not demonstrably false by (a).
- (e) Therefore the harm should be taken seriously and is often large enough to outweigh benefits of the research.

(a) Religion makes untestable claims about harms and benefits.

Religion makes claims about metaphysical facts that are generally not empirically verifiable. Most religions tend to make claims about the meaning of life, the terms of any pre-life experience or after-death experience, and the absolute morality of particular human actions. Many humans —

particularly religious humans — find these questions to be of great importance, even though they are not typically susceptible to empirical verification.

Even when religious traditions generate falsifiable empirical claims, adherents often adopt methods of empirical investigation that allow for religion to retain meaning even if truth claims are, in a particular instance, apparently false. For example, religious groups whose charismatic leaders have made testable prophecies tend to remain faithful, even when those prophecies do not occur as the leader or group expected (Festinger, Riecken and Schachter, 1956). Mormons argue that spiritual knowledge must be ultimately obtained through feelings of divine experience rather than through scientific study; some Mormon historians have become increasingly postmodern in their approaches to understanding aspects of Mormon history that seem to be in tension with theological claims of the church (Duffy, 2008). The Kamajor secret society in Sierra Leone claimed to endow the power to withstand bullets following initiation and continued adherence to religious principles (Kelsall, 2009, 129-137).² When adherents observed examples of society members succumbing to bullets, they believed that the individuals were not sufficiently scrupulous in their adherence to the requirements of the secret society such as abstaining from sex and not eating pumpkin. All this is to say that many individuals have demonstrated that they find religious beliefs to be important and that they hold these religious beliefs to be largely outside of the realm of empirical verification.

(b) People with different beliefs can view identical treatments as clear harm or clear benefit.

Because of the incompatibility of competing religious claims, people with different beliefs can view identical change manipulations as clear harm and clear benefit. Any treatment that changes religiosity could be seen as a clear benefit by people whose religious ideal points are closer to the subject's post-treatment religiosity, while the same change will be seen as a clear harm by people with religious ideal points that are closer to the subject's pre-treatment religiosity.

²Thanks to Rebecca Nielsen for bringing this example to my attention.

(c) All treatments that meaningfully change religiosity will be a clear harm to someone.

Because of the exclusivity of many religious claims, there are proponents and opponents of almost all religious positions. Atheism, Bhuddism, Christianity, Hinduism, and Islam (to name a few religious traditions in alphabetical order) are largely seen as incompatible and mutually exclusive by most adherents and detractors. Thus for all meaningful religious changes that could be encouraged via treatment, there are some individuals who would view the change as religiously beneficial and some who would view it as harmful.

(d) Such a claim of harm is not demonstrably false by (a).

These claims of harm cannot be tested because many of them rely on untestable claims about an afterlife. Many religions claim that eternal punishments exist in an afterlife for individuals who do not practice religion appropriately. If these claims are true, such infinite harm would certainly outweigh physically observable benefits following a particular course of action.

(e) Therefore the harm should be taken seriously and is often large enough to outweigh benefits of the research.

Without an ability to adjudicate between true and false claims of harm, social scientists that use meaningful change manipulations will be causing harm from the subjective view of the participant. If the perceived harms are substantial, then the potential harm to subjects is likely to outweigh the benefits of the research in many circumstances. At very least, experiments that attempt to change religiosity should explain why the potential metaphysical harms are small relative to the benefits of the research.

Scholars should be wary of their own judgments about which harms should be taken seriously. Natural scientists are more likely than the average individual to be agnostic or atheist (Larson and Witham, 1998), and presumably social scientists follow suit. If academics are inclined against religion, they may tend to consider heavily the benefits of reduced religiosity and discount the harms. On the other hand, religious adherents are likely to feel the opposite: that the harm of becoming more religious is negligible relative to the harm of becoming less religious. Fundamentally, these

individuals may simply not agree on the definition of a “harm.” While damnation at the hands of a wrathful God seems like serious harm to some, contributing time, money, and energy to an organization with false claims about deity and discriminatory practices may seem like more serious harm to others. Because the truth claims of religious are not susceptible to empirical verification, this debate will not be resolved. This suggests that as scholars, we need to carefully consider harms that may be obvious to adherents but are not obvious to us because of our biases. Without an absolute standard for judging with change manipulations are ethical, there is strong reason to adopt the conservative position that scientific researchers should not do experiments involving change manipulations of religiosity without substantial justification.

Hypothetical Examples

To illustrate why the ambiguity of religious harms and benefits makes change manipulations problematic, I consider a series of potential experiments and how religious individuals might react to them. I discussed these thought experiments with three Mormons and I note their reactions to each, though they are hardly a representative sample of any religious population.

Imagine a study in which researchers at a national atheist group would like to improve the effectiveness of its materials for persuading Mormons. The group identifies a sample of Mormons and sends them mailings with elements of the mailings randomized. One year later, the researchers follow up with these individuals to measure the number of individuals who have decreased faith in the teachings of the Mormon church or left Mormonism. Perhaps not surprisingly, all three Mormons I talked with found this study ethically troubling.

However, consider an essentially identical study design: researchers in the Public Relations department of the Mormon church would like to improve the effectiveness of its materials for persuading non-Mormons. The group identifies a sample of non-Mormons and sends them mailings with elements of the mailings randomized. One year later, the researchers follow up with these individuals to measure the number of individuals who have increased faith in the teachings of the Mormon church or have joined Mormonism. At this point, the three Mormons I interviewed typically said, “That sounds fine to me,” and then a few moments later, “I see where you are going

here.” I then invited each of them to imagine that they were an atheist ex-Mormon who believed that the church was harmful and discriminatory rather than a practicing member of the Mormon church. Each agreed that their evaluation of the last two thought experiments would probably flip, with the deconversion experiment seeming ethical and the conversion experiment seeming unethical.

Although this example hinges explicitly on untestable metaphysical claims, experimental designs that present verifiable facts in an attempt to change religiosity are similarly problematic because they implicitly involve these same claims. For example, it might be possible to decrease the religiosity of Mormons by presenting them with particular pieces of true information about Mormon church history that do not square with the somewhat sanitized history that is taught within the church. Many former Mormon adherents credit their loss of faith to the experience of receiving such information.³ However, it would probably be unethical to carry out an experiment that randomized exposure of this information to self-identified Mormons. Even if the historical information were entirely accurate, the resulting changes in religiosity are harmful from within metaphysical framework of Mormonism and this potential harm should not be risked without substantial justification.

Now, I turn to hypothetical measurement manipulations and find that these are less problematic. Imagine a study in which university researchers identify a sample of religious individuals and a similar sample of non-religious individuals. The individuals in the study are randomly paired, sometimes with a co-religionist and sometimes not. The individuals play a dictator game to assess the generosity of one of the partners in each pairing. There is no randomization of religion. The manipulation is entirely to measure a particular outcome — altruism or cooperation — for each pairing. This strikes me as unproblematic, and the three Mormons that I discussed this example with agreed.

Adding an element of deception to a hypothetical measurement manipulation induced some concerns about ethics. Consider an experiment in which university researchers participate in online conversations with Mormons and act as if they are interested in learning about the beliefs of these Mormons. They request information about the Mormon church and invite the Mormon to share their personal feelings of belief (called a “testimony” in Mormonism). Each request contains some

³http://www.whymormonsquestion.org/wp-content/uploads/2012/04/Survey-Results_Understanding-Mormon-Disbelief-Mar2012-1.pdf.

framing element that is randomized. For example, we could imagine that the name of the requester is randomized to be a name that is commonly associated with a certain ethnicity. The researchers then study differences in response rates and the content of the information provided by the Mormons in the study.

The three Mormons that I discussed this with had the gut reaction that this was unethical. I then asked them to imagine that the element of deception could be largely removed: that the Mormons would be adequately debriefed following the study. This led each of the Mormons I discussed this with to decide that the study was probably ethical. One of them — an academic economist — said “I wouldn’t really want to do an experiment like this myself because of the amount of deception, but in other domains we tolerate this kind of minimal deception in experiments all the time.” I argue that this experiment would be ethical from the standpoint of religion because it is primarily measuring religious response rather than changing religiosity. The ethical issues are issues of deception and informed consent rather than religion per se.

2.3 An Argument from the Impossibility of Informed Consent

- (a) People have a right to determine their own religious beliefs.
- (b) This requires informed consent of potential consequences of treatment.
- (c) Few individuals will consent to experimental manipulation of their religiosity and those that do are unlikely to be representative.
- (d) Therefore we cannot experimentally learn about the effects of religion in most populations.

(a) People have a right to determine their own religious beliefs.

Absent other considerations, individuals should have autonomy to choose their religious beliefs. Although these beliefs are often unverifiable, for adherents, the consequences of these beliefs are critical. In some religious traditions, individuals believe that they are determining the fate of their own eternal identity with these beliefs. In practice, no individuals are truly “free” to choose religious beliefs outside of the context from which they come. The strongest determinant of an individual’s religious tradition is the religious tradition of their parents (Spilka et al., 2003, 115-116), suggesting that socialization and family dynamics constrain religious choice. In open societies,

individuals face constant influences from peers to modify, change, deepen, strengthen, or relinquish religious beliefs. In general, such influences are morally defensible on the grounds that individuals should have freedom to associate (or disassociate) and communicate freely about matters they feel are important.

Social scientists as a class do not have the same rights to engage in such conversations. The fact that individuals are influenced by external forces when choosing their religious beliefs does not give social scientists the right to manipulate individuals' religious beliefs. While a social scientist as a private individual has rights to attempt to influence the religiosity of friend, the same social scientist does not have the inherent right to attempt to "set" the religiosity of the same friend to a particular level for scientific purposes without consent.

(b) This requires informed consent of potential consequences of treatment.

I argue that informed consent is necessary for any experiment that intends to experimentally change religiosity because there is potential for harm to an individual who has their religious beliefs or experience manipulated for the sake of scientific inquiry. In Section 2.2, I considered metaphysical harms that cannot be definitively proven. Here, I consider physical and emotional harms for which evidence is readily available.

On a basic level, there may be harm to the subject if an experiment causes the subject's future self to adopt a level of religiosity that their current self would not like. I argue that this is true, *even if the subject's future self is satisfied with the new level of religiosity that they are assigned by the researcher*. This is because the right of an individual to choose their religious beliefs includes the right to fully determine one's own religious beliefs in the future. Such a change would generally only be ethical if participants were fully informed that their level of religiosity might be changed if they participated in the experiment and that they might or might not be happy with their assigned level of religiosity.

Individuals can often face harm from becoming more or less religious and they should generally free to choose such harm for themselves. Examples of harm from decreased religiosity are relatively easy to list. For example, individuals that disassociate with the Jehovah's Witnesses may be

shunned by family and friends. Likewise, Muslims who convert to other religions may face a capital sentence for apostasy under some interpretations of Islamic law. Similar social costs are imposed on apostates from a wide variety of faiths. More broadly, individuals who become less religious may experience a painful loss of community, a feeling of disorientation, alienation from social connections, mental illness, and other costs. There may also be benefits to decreased religiosity that potentially counterbalance these costs. Individuals may be controlled or manipulated within religious organizations, and may face sexual, physical, or emotional abuse as a result of affiliation with religious organizations. There is also potential harm in becoming more religious. Individuals have faced substantial religious persecution throughout history and still face human rights violations in some parts of the world. Individuals may also face mental illness as a result of increased religiosity, including tendencies toward obsessive-compulsive disorders and scrupulosity (Spilka et al., 2003, 511). Certainly, these harms do not happen to everyone who changes religiosity, but researchers cannot know *ex ante* which individuals will be harmed so these should be serious considerations even if only a small portion of individuals are affected. At a minimum, the potential for these harms will generally require that experimental subjects give informed consent.

(c) Few individuals will consent to experimental manipulation of their religiosity and those that do are unlikely to be representative.

Experiments on religion that use informed consent will only have subjects who knowingly agree to the possibility that a research will attempt to substantially change their religiosity based on randomized assignment to a treatment condition. What sorts of subjects will opt into such an experiment? The most obvious set are those whose attachment to their current and future religious identities is so weak that they are willing to allow their religiosity to be manipulated. I expect that there will be relatively few such individuals because most people seem unwilling to casually change their religious beliefs. A second set of individuals may believe that they are not susceptible to manipulation of their religiosity and agree to participate, but I believe these will also be relatively rare.

Neither of these types of individuals are representative of the adherents to a particular religious

tradition or the broader population. It seems unlikely that most meaningful questions about the role of religion in politics could be answered using a sample of people who are willing to subject their religiosity to researcher manipulation, either because they are apathetic or extremely self-confident. After all, what does it mean to say that religion causes some outcome among a subgroup of people who are willing to determine their religiosity or religious beliefs according to a coin flip.

(d) Therefore we cannot experimentally learn about the effects of religion in most populations.

As a result, the necessity of informed consent means that even when manipulation of religion for experimental purposes is ethically defensible, it is likely to be impossible. Scholars will usually not be able to experiment on the population to which their theories apply, but rather the self-selected subset who would give consent. Perhaps there are some questions for which such a sample might be able to adjudicate between competing theories, but such cases seem unlikely. In general, social scientists care about the effects of religion precisely because it is a powerful, consequential force in the lives of individuals. Experiments on the small subset of individuals for whom religion is so inconsequential that they will allow it to be randomized are unlikely to be informative.

When are experiments that change religiosity ethical?

Experiments that change religiosity may be ethical if researchers can show that there is direct benefit that comes to subjects or others from changing a religious belief or practice that outweighs untestable claim of harm. For example, militant Jihadists may believe that suicide bombings are a religiously encouraged activity, but research that discourages suicide bombing by manipulating religiosity will bring enough demonstrable benefit to both the subject and potential bombing victims that it is ethical (and perhaps imperative) to carry out the change manipulation. Even in this instance, much relies on the assumption that the researcher can apply a treatment that will only have ameliorative effects on radicalization. This is a problematic assumption. On one hand, if the assumption holds, then the experiment is unnecessary because we already know that the treatment will have deradicalizing effects. The experiment should be scrapped and a full-scale intervention

should be rolled out. On the other hand, if we are not sure enough about the sanguine effects of the treatment, then we are unlikely to be sure enough that we are not causing harm to carry out the experiment. The benefits in such a setting are still likely to outweigh the risk, but scholars should be cautious.

Manipulations that are similar to circumstances in the daily lives of participants may be ethical because they pose no more risk to religiosity than the risk the subject faces from daily influences. For example, Gervais and Norenzayan (2012) had some subjects complete tasks that required analytical thinking and found that this decreased religious belief in the short term. In general, I argue that decreasing religiosity experimentally is problematic, but the tasks involved were to spend time looking at a picture of Rodin's statue *The Thinker* (control condition: a picture of the *Discobolus* of Myron) and a verbal fluency task in which treated subjects were given words such as "analyze, reason, ponder, think, rational" while control subjects were given words such as "hammer, shoes, jump, retrace, brown, etc." Given that these seem like relatively innocuous treatments, the experiment does not seem problematic despite the finding that these tasks temporarily reduced religiosity.

Are manipulations that change non-religious beliefs ethical?

Religious beliefs are not the only kind of belief that social scientists might want to manipulate experimentally. For example, researchers might want to know whether making someone more politically liberal affects their political behavior, or whether assignment to a civic education program increases belief that democracy is the best system of government. Do researchers face the ethical challenges outlined above when manipulating non-religious beliefs?

I argue that the differences between religious beliefs and non-religious beliefs are best viewed as differences of degree rather than differences of kind. The premise of my argument in Section 2.2 is that religious beliefs are generally not empirically testable. Neither are many political beliefs. The premise of the argument in Section 2.3 is that people have a right to determine their religious beliefs, in part because those beliefs have physical costs and consequences. Political beliefs also have costs and consequences, so informed consent is necessary and the subset of subjects willing to

subject their political convictions to a coin flip may be small and unrepresentative.

It is primarily on the grounds of the argument in Section 2.1 that we can start to distinguish political and religious beliefs. Arguably, most political beliefs have less drastic metaphysical consequences than religious beliefs. Few people believe that they would face eternal punishment for changing their political beliefs than for changing their religious beliefs. This lessens the potential for harm considered in argument 2.2, meaning that experiments manipulating political beliefs have fewer costs to weigh against the benefits, so perhaps more experiments will be possible.

It is also possible that argument 2.2 may also provide some distinction between religious and political beliefs if some political beliefs are also more susceptible to evidence than religious beliefs. Political orientations tend to make at least some claims that are verifiable and these claims can be linked back to philosophically defensible aspirations of individuals to autonomy, health, happiness, etc. Perhaps it is ethical to change political beliefs in ways that unambiguously expand the scope of human freedom and happiness.

Nevertheless, my argument raises serious questions about interventions intended to promote a variety of beliefs. If it is unethical to administer a treatment that turns a subject into an apostate who eventually faces death at the hands of their former religious community, then it seems similarly unethical to administer a treatment that turns a subject into a democracy activist who eventually faces imprisonment in a dictatorial regime, even if we believe that spreading democratic norms is good. These are extreme examples, but even when the potential consequences are far smaller, subjects should be warned about the costs and benefits of manipulations designed to change their beliefs. Many of them may opt out, perhaps making the experimental results unrepresentative to the point of uselessness.

3 Examples

The previous sections distinguished measurement manipulations from change manipulations and argued that the latter are likely to be more ethically problematic. In this section, I discuss examples of experiments that have been proposed or published in light of the principles I have outlined. The first purpose of this is to see whether my framework offers useful guidance for considering

the ethics of experimental manipulation in religious settings. The second purpose is to evaluate whether applying my framework to existing studies leads to judgments that are surprising or seem wrong.

I first consider a failed experiment proposal of my own which led to the arguments I have developed here. I then discuss other examples of experimental work.

A Failed Experiment Proposal

As part of a conference for “Experiments in International Relations” held in Park City, September 21-22, 2012, I prepared and presented an experiment proposal entitled “Why do Muslim clerics issue *fatwas* supporting militant Jihad?” The purpose of the proposed experiment was to test arguments I have made in other work⁴ that Muslim clerics are sensitive to strategic considerations when they decide to express more or less extremist rulings in their Islamic legal rulings (fatwas).

I proposed to first identify a list of clerics who issue Islamic legal rulings over the Internet in response to questions from lay Muslims. I would then request fatwas from these clerics, hiding my identity so that the cleric would believe that they were simply answering yet another of the tens and hundreds of fatwa requests they answer daily. These fatwa requests would reference some current event and then ask a for a ruling that had potential to be Jihadist or not. Specifically, I planned to ask clerics three questions on (1) the permissibility of participating in democracy (in the wake of the Arab spring), (2) whether it is permissible to participate in the Syrian uprising and whether it is a legitimate Jihad, and (3) whether suicide operations are permissible in the Syrian uprising. I would randomize the addition of several statements to the baseline fatwa request, intended to make clerics think of either the costs or benefits of issuing certain types of rulings. To cue the costs of being Jihadist, I would add “I know that in the past, some clerics have faced punishment or been imprisoned for speaking freely on issues like this one.” To cue the credibility benefits that I theorize accrue to Jihadists, I planned to add “In the past, I and many of the other brothers have found that those who are willing to speak freely on this issue are the most trust-worthy of the clerics.” Finally, to test whether the rulings of other clerics matter, I considered including “I have read the

⁴See Nielsen (2013), accessible at <http://dash.harvard.edu/handle/1/11124850>

ruling by [CLERIC NAME] which rules on this issue as follows: [RULING]. What is your opinion on this matter?” These would have been compared to a fourth, baseline condition with just the fatwa request.

I then planned to collect responses from clerics and look at two outcomes: response rate and the content of the fatwas for those clerics that responded with fatwas. My theory would have found strong support if a reminder of the costs made clerics’ rulings more democratically oriented and less prone to support violent activism of various kinds, while a reminder of reputation benefits of Jihadism would have the opposite effects.

Not surprisingly, the bulk of the comments I received about this proposal related to ethical concerns.

1. What will an IRB say? And if an IRB approves, is it still unethical?
2. Can we make this more ethical by only using de-radicalizing treatments?
3. Could you openly invite the clerics to participate so that there is no deception? Is informed consent through email a problem?
4. Is there a possibility that anyone could be harmed or killed as a result of this research? Is there any way to justify the commissioning of texts that advocate violence? Commissioning potentially violent fatwas is a bad idea.
5. Don’t do this experiment because it could backfire publicly and threaten the entire experimental social science enterprise!
6. Should we be randomizing other people’s religious beliefs?

In my proposal, I had anticipated (and invited) virtually all of these critiques except the last — should scientists be randomizing other people’s religious beliefs? I agreed that something seemed problematic with randomizing religious beliefs, and as a person with long-standing experience with religion, I felt sensitive to such considerations. On the other hand, this had not occurred to me, while basically all of the other critiques had. I spent the remainder of the conference serving as something of an ethical baseline (Discussant of another paper: “I find this ethically problematic, but not as problematic as that fatwa paper.”), which naturally forced me to think seriously about both justifying and abandoning the project. This chapter is the result of my attempt to respond to the critique I had not anticipated: that an experiment involving religion was inherently problematic.

As I conceptualized it, my experimental proposal to elicit fatwas was essentially a measurement manipulation. I did not explicitly conceive of it in these terms — I had not yet developed either the terms or the concepts — but I believe this is why I was blind-sided by the criticism that I was randomizing religious beliefs. In my other work on Islamic fatwas, I use fatwas as a vehicle for measuring cleric ideology (Nielsen, 2013). In doing so, I work from the basic assumption that their ideology is fixed at the time of writing, and that writing is an observable manifestation of ideology. I have not generally considered the possibility that the process of fatwa-giving *changes* a cleric’s ideology, at least in the short term. Rather, the cleric already has an ideology that is latent until they are asked about a particular situation. Thus, I believed that fatwas could be used as a very authentic form of survey, and that my framing manipulations were akin to a survey experiment manipulation intended to explore different aspects of a fixed ideology rather than to actually change the religious ideology of the cleric in any way.

Most of the ethical concerns raised by participants at the conference seem to have been premised on the belief that my intervention might in be a change manipulation, either for the clerics themselves, or for their followers. It is possible that by asking clerics about violence and Jihadism, I might force them to think about issues more deeply and that they would come to more radical positions than they previously held. The framing experiment is certainly an attempt to temporarily focus the subject on a particular aspect of a decision; this itself could be viewed as a way of changing the subject’s religiosity, depending on how one interprets the psychology of framing. If fatwas are effective at changing the opinions and practices of lay Muslims, then generating additional fatwas from artificial questions could actually distort “natural” Muslim orthodoxy and orthopraxy. Ethical concerns about the unintended consequences of eliciting texts that might justify violence are also founded partially on the idea that the experiment would be changing religious behavior rather than merely measuring it. Those concerned about deception were at least partly worried that pretending to be a co-religionist with a sincere request would be insensitive to the religious context and subtly change the religious experience of the cleric or lay Muslims.

I believe that there could be ways to eliminate the risks of violence and issues with deception and informed consent. If these other ethical issues were resolved, would it be ethical for a non-Muslim

to participate in the practice of fatwa-asking and fatwa-giving for the purposes of social science research? I think this is an open question that hinges on whether I would be simply measuring cleric ideology, or whether the religious beliefs of clerics or lay Muslims would change as a result of the experiment. Put differently, would it bother clerics and followers to find out that a particular ruling was actually in response to the question of a social science researcher who had randomized parts of the question rather than a sincere religious seeker. My hunch is that this would actually not change clerics' or lay Muslims interaction with the text (in other words, it would not be blasphemous to do this experiment), but this is an open question.

Having described the specific project that prompted me to develop the ideas in this chapter, I now apply my framework to a series of studies to illustrate the usefulness of my distinction between measurement manipulations and change manipulations in religious contexts.

Measurement Manipulations

My framework suggests that experiments that primarily measure religiosity rather than changing religiosity to estimate effects are more likely to be ethical. Is this really so when I examine actual studies that appear controversial at first blush?

A seemingly hard case is the paper “The Economics of Faith: Using an Apocalyptic Prophecy to Elicit Religious Beliefs in the Field” (Augenblick et al., N.d.).⁵ In this paper, the authors examine a small religious group following Harold Camping who prophesied that the end of the world would occur on May 21, 2011. The authors test whether “beliefs in the prophecy among Family Radio members [Camping’s followers] are a matter of external profession rather than inner conviction,” by offering them financial incentives of either \$5 prior to May 21st, or varying amounts up to \$500 after May 21st. The amounts and probability of payment were experimentally manipulated to allow estimation of the discount rate of participants. As a comparison group, the authors gave the same choices to a sample of Seventh Day Adventists who similarly believe that the end of the world is “imminent” but did not believe that the end would come on May 21st.

The authors summarize their findings:

⁵ Available at <http://faculty.haas.berkeley.edu/ned/WaitingForTheEnd.pdf>

The evidence indicates that the vast majority of Family Radio members held extreme beliefs even in the face of direct financial costs. nearly all Family Radio subjects preferred \$5 dollars today to any amount up to \$500 payable after the Rapture, regardless of the probability of implementation. At the same time, the SDA members made choices consistent with time preference parameters estimated in laboratory studies (Frederick, Loewenstein, and O'Donoghue 2002). Taken together, these findings indicate that the Family Radio members held sincere and full beliefs in the prophecy, with little apparent elasticity. This finding underscores the role of sincere faith in the demand side of religion, and simultaneously rules out two alternatives: evidence-based beliefs, and faith-related activities that occur exclusively due to social factors.

I argue that my framework correctly illuminates why this study is ethical (at least with respect to issues of religion). Although the paper uses the word “experiment” over 80 times in reference to the research, the experimental portion of the research is simply an innovative way to measure the discount factors of two religious groups. There is nothing experimental about the actual assignment of religious beliefs. This means that the portion of the paper that essentially estimates the effect of a change in religious belief about the Camping prophecy is actually observational rather than experimental. Thus, in the framework I introduce, this paper is close to a pure measurement manipulation. Contrary to what I might have assumed from the title and abstract, I do not think this paper exploits a vulnerable religious group or attempts to tinker with “odd” religious beliefs for the purposes of social science.

Similar principles apply to a genre of studies that use experimental manipulations to measure cooperation between individuals and then test whether the religiosity of participants conditions their level of cooperation (Ruffle and Sosis, 2007; Tan and Vogel, 2008).⁶

Weak Change Manipulations

The other major category of existing studies are what I would call “weak change manipulations.” These are typically framing experiments in which the experimenter attempts to cue religion by exposing the subject to text or images that will remind them of their religious identity rather than some other identity. I call this weak change manipulation because the research is attempting to manipulate the religiosity of the respondent by making them temporarily more religious. However,

⁶Ruffle and Sosis (2007) is available at <http://www.bgu.ac.il/~bradley/Publications/paypray.pdf>. Tan and Vogel (2008) is available at http://www.econstor.eu/bitstream/10419/23811/1/240_Tan_Vogel.pdf.

in most of these studies, the authors appear to assume that priming is merely activating an already existing religious identity, rather than fundamentally changing the long-term religiosity of a person. Based on my framework, I argue that these experiments need more justification, but are likely to be ethically defensible in most cases.

The major weakness of these studies is that they tend to overclaim when interpreting their results. For example, a recent working paper that uses a priming experiment is titled “The effects of religion on social cooperation: Results from a field experiment in Ghana” (Parra, N.d.).⁷ However, this experiment cannot identify the effects of religiosity unless it is meaningfully making respondents more religious. The treatment does not seem to do that. Rather, participants are primed with religious imagery — a picture of a Christian choir, a Muslim praying, the Ka’ba in Mecca, and a crucifix and prayer beads — and then participate in a dictator game to (experimentally) measure social cooperation. This priming intervention seems too weak to support claims about the “effects of religion.” It is stretch to extrapolate from the momentary burst of religiosity after viewing a religious image to estimate what would happen if individuals became more devout in the long-term. Under this interpretation, the experiment is likely to ethical but limited in its relevance to understanding the social world. Alternatively, if the experimenter believes that such priming *is* inducing substantial or long-lasting in religiosity, then the ethics of the experiment should be questioned.

Experiments that prime religious identities are quite common, suggesting that scholars and review boards find them ethical.

Strong Change Manipulations

I could find not published examples of strong change manipulations — experiments where manipulations were intended to change participants religious beliefs in a substantial and long-term way. This lack of examples probably supports my assertion that such experiments are typically unethical. It could be the case that finding opportunities to meaningfully change subjects religious beliefs is more difficult than other types of experiments on religion, but I suspect that some opportunities

⁷Parra (N.d.) is available at <http://www12.georgetown.edu/students/jcp29/ghanaexp3.pdf>.

are available. Scholars appear not to have taken those opportunities, suggesting the existence of a taboo.

4 Conclusion

In this paper, I have argued for a distinction between measurement manipulations and change manipulations in experiments involving religion. While neither ideal type is possible, the continuum from measurement to change is an important dimension for evaluating the ethics of experimental manipulation of religion. Manipulations that primarily measure the religiosity of an individual without changing it will generally be ethical, provided that they comply with other ethical criteria. Experiments that change the religiosity of participants may be less ethical. A middle ground exists with experiments that prime religious identities because it is not clear whether the manipulation is changing a subject's religiosity or merely increasing the salience of a stable religious identity without changing it. This probably hinges on a careful definition of religion and religiosity that I have failed to provide here.⁸ My argument does point to a dilemma with these studies: either priming interventions are not as informative about the effects of religiosity as some proponents suggest, or they are possibly unethical.

I do not claim that the spectrum from pure measurement manipulations to pure change manipulations is the only dimension for ethical evaluation of experiments in religious contexts. I do think the distinction I introduce is instructive for thinking about the ethics of some experiments that are being carried out in the contemporary social sciences, but other distinctions may also be helpful.

There is no empirical way to establish that the arguments I advance here are the right way to think about the ethics of randomizing religion. I believe it supports my case that the types of experiments that I argue are clearly unethical are the same types of experiments that seem largely non-existent on the scholarly literature. I interpret this absence as partly the effect of a taboo, but it may have other causes. I am aware of some experiments in the planning stages that I would classify as strong change manipulations. If these experiments proceed, the reaction to them will

⁸Defining "religion" is not trivial. See Platvoet and Molendijk (1999).

be instructive as to whether the ethical claims I make here resonate with social scientists more broadly.

However, we should be careful that we do not exclusively benchmark against the ethical beliefs of social scientists. Religious individuals and communities have important stakes in this debate and scientists are unlikely to automatically consider or represent these other views. One way to explore the ethics of religious experimentation would be to survey individuals outside of academia to find out the types of experiments that they would consider ethical. In fact, one could propose an (ethical?) experiment on the ethics of experiments on religion in which respondents of differing religiosity are presented with a brief summary of an experimental design where key parts of the description were randomized. In order to test my argument, such an experiment would need to vary the degree to which the religiosity of respondents would be manipulated to see whether religious and scholarly communities agree that change manipulations are less ethical. Note that this survey experiment would be ethical under my proposed framework, but might not be ethical under others.

My argument implies that it may be difficult to ethically learn about the effects of individual religiosity from experiments. The only way to experimentally test the effects of changes in religiosity is to change subjects religiosity arbitrarily. Otherwise, subjects' observed levels of religiosity are endogenous to their own choices and are likely to be related to the outcome of interest through confounding variables. Experiments are not the only way to learn about causal relationships, but removing experiments from the toolkit of social scientists studying religion could be a real impediment to progress. Still, scholars cannot do unethical things simply because we would like to know more about the social world.

My fear is that these challenges will deter young scholars from working on issues of religion and politics. If experiments are privileged and professional rewards accrue to those who use them, young political scientists may naturally gravitate toward topics and questions for which conducting experiments is not fraught with ethical complications. Given the importance of religion in politics, this would be a real loss. My remedy is to propose that scholars should entertain unethical research designs — not in order to implement them (please do not!) — but because pondering an unethical design can often lead to ideas about how it could be made ethical. If scholars are too quick to

reject some approach as unethical, a breakthrough that allows the research to proceed will never occur. This work may be more difficult, but ultimately the importance of understanding the role of religion in politics demands that we face the inherent ethical hurdles head on rather than diverting our best efforts to more tractable topics.

References

- Augenblick, Ned, Jesse M. Cunha, Ernesto Dal Bo and Justin M. Rao. N.d. "The Economics of Faith: Using an Apocalyptic Prophecy to Elicit Religious Beliefs in the Field." Available at <http://faculty.haas.berkeley.edu/ned/WaitingForTheEnd.pdf>, accessed 23 April, 2013.
- Batson, C. Daniel. 1977. "Experimentation in Psychology of Religion: An Impossible Dream." *Journal for the Scientific Study of Religion* 16(4):413–418.
- Batson, C. Daniel. 1979. "Experimentation in Psychology of Religion: Living with or in a Dream?" *Journal for the Scientific Study of Religion* 18(1):90–93.
- Duffy, John-Charles. 2008. "Can Deconstruction Save the Day? "Faithful Scholarship" and the Uses of Postmodernism." *Dialogue: A Journal of Mormon Thought* 41(1):1–33.
- Festinger, Leon, Henry W. Riecken and Stanley Schachter. 1956. *When Prophecy Fails*. University of Minnesota Press.
- Gervais, Will M. and Ara Norenzayan. 2012. "Analytic Thinking Promotes Religious Disbelief." *Science* 336 (6080):493–496.
- Kelsall, Tim. 2009. *Culture Under Cross-Examination: International Justice and the Special Court for Sierra Leone*. New York: Cambridge University Press.
- Kuklinski, James H., Michael D. Cobb and Martin Gilens. 1997. "Racial Attitudes and the "New South"." *The Journal of Politics* 59(2):323–349.
- Larson, Edward J. and Larry Witham. 1998. "Leading scientists still reject God." *Nature* 394:313.
- Nielsen, Richard A. 2013. *The Lonely Jihadist: Weak Networks and the Radicalization of Muslim Clerics* PhD thesis Harvard University.
- Parra, Juan Carlos. N.d. "The effects of religion on social cooperation: Results from a field experiment in Ghana." Accessed at <http://www12.georgetown.edu/students/jcp29/ghanaexp3.pdf> on 23 April, 2013.
- Platvoet, Jan G. and Arie Leendert Molendijk, eds. 1999. *The Pragmatics of Defining Religion: Contexts, Concepts, and Contests*. Leiden: Brill.
- Ruffle, Bradley J. and Richard Sosis. 2007. "Does It Pay To Pray? Costly Ritual and Cooperation." *The B.E. Journal of Economic Analysis & Policy* 7 (1):18.
- Spilka, Bernard, JR Ralph W. Hood, Bruce Hunsberger and Richard Gorsuch. 2003. *The Psychology of Religion: An Empirical Approach*. Third edition ed. New York: The Guilford Press.
- Tan, Jonathan H.W. and Claudia Vogel. 2008. "Religion and trust: An experimental study." *Journal of Economic Psychology* 29 (6):832848.
- Yeatts, John R. and William Asher. 1979. "Can We Afford Not to Do True Experiments in Psychology of Religion? A Reply to Batson." *Journal for the Scientific Study of Religion* 18(1):86–89.